OPTIONAL FORM NO. 10
MAY 1962 EDITION
GSA FPMR (41 CFR) 101-11.6

UNITED STATES GOVERNMENT

Memorandum

STAT

TO   :   C9
          CØ3

DATE: 22 July 1969

FROM :   C95

SUBJECT: Information Storage and Retrieval

   1. Attached is our recommendations on Information Storage and Retrieval as related to the COINS network and to the Agency.
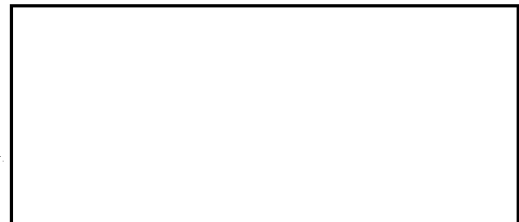
STAT

Incl: a/s

Buy U.S. Savings Bonds Regularly on the Payroll Savings Plan

INFORMATION STORAGE AND RETRIEVAL

RECOMMENDATIONS

22 July 1969

STAT

Information Storage and Retrieval

Introduction

From 27 November 1968 until the present, C95 has been studying both the multiplicity of user languages in the COINS network and the problems and discipline surrounding Information Storage and Retrieval (IS&R) in general. We have examined in detail, and from several viewpoints, the three existing COINS languages: TILE, TORQUE, and DIAOLS; an Agency query facility, SPECOL; a commercially available system, TDMS; and an IS&R system designed and documented under an R45 contract, QUL. A technical report on our findings has been published as a memo dated 29 May 1969. This report concerns itself with our recommendations regarding both the COINS problem of several user languages and NSA's problem of developing a comprehensive IS&R facility.

Requirements

We feel that any IS&R system should place in the hands of a wide spectrum of analysts a software tool to provide the rapid and timely formulation of and response to actions addressed to data files. There are three important parts in the previous statement. First, the phrase "of a wide spectrum of analysts" implies a community of users with diverse interests and capabilities. Second, "rapid and timely formulation" implies a language of a natural form, designed for non-computer professionals which can be presented in a stratified manner without restricting anyone to a limited level of capability. Third, "rapid and timely response" implies both a high degree of man-machine Interaction from remote consoles and non-sequential access to on-line files of a variety of structures. Consequently any IS&R system should be totally integrated, well balanced, and have a strong potential for growth.

Possible Solutions

In particular regard to the COINS problem of multiple user languages, several solutions have been proposed in the past. These can be briefly summarized as (1) one language translatable to each existing language, (2) all existing languages translatable to all other existing languages, and (3) all existing languages translatable to one internal language. All of these solutions, although politically expedient, either are technically unsound (since the meanings of queries in a given system are totally dependent upon its permissible file structures) or require unwarranted and extensive gimmickry.

Another solution would be to select the most capable of the existing systems and implement it as the standard network facility. In this case our choice would be DIAOLS (NIPS). This system is favored

over TILE and TORQUE because (1) it is simple in structure and use; (2)
it has the most accommodating filing system of the existing facilities;
(3) it is operational and there is a relatively extensive body of
experience with its use; (4) it is a quite complete system relative to
the other existing systems; and (5) if chosen, it would be the easiest
to install since it already exists on the IBM-360.

25X1   In our case DIAOLS (NIPS) could exist on an IBM-360 and presumably   25X1
use the UNIVAC-494 as a switch.)

Certainly there are definite consequences to such a solution. The
most important ones are (1) the major file conversions that would be
necessary, (2) the alterations that might be desirable to expand the
capabilities of the query language, (3) the fact that the system design
and filing structure severely limits the achievement of immediate user
responses and on-line File Maintenance procedures, (4) the absence of
a fully integrated system from the user's standpont, and (5) the im-
possibility of attaining any significant degree of sophistication in
the filing structure would result in correspondingly low level queries.

If one looks beyond the presently available systems, another
solution would be to select and implement a facility such as QUL. In
its present state, it is most readily shaped to any particular needs,
and in general, it exceeds the capabilities of all existing COINS
facilities. It is well-balanced and fully integrated (as our technical
report illustrates) and the system admits of a very high degree of man-
machine interaction.

Needless to say, we have a certain amount of concern over QUL. To
begin with this system does not now exist on any computer and it would
be at least a two year endeavor to implement QUL for some stand-alone
system. Moreover, because of how little is known at this time about
computer networks, the existence of a new retrieval system on any one
computer is not necessarily equivalent to its existence in network.
The major reason for QUL's power and flexibility is the unique file
inversion technique employed. Still, there exist no disciplines for
these types of inverted files as they apply to the kinds and amounts
of data with which we must be concerned.

Recommendations

It is obvious to us that there is no simple, expedient, or
painless solution for either the problems of COINS or the problems of
developing an in-house IS&R facility. We recommend the following:

1. If it is desirable to stress the immediate development of the
network then we feel the best course of action is to select DIAOLS
(NIPS) as the system language, which implies the selection of IBM 360
as the system hardware. In addition to what has already been said

about DIAOLS, there would be some additional benefits: (1) The
three agencies could begin work on achieving hardware com-
patibility which will probably have to be done regardless of which
solution is selected, and (2) the value of the COINS network inde-
pendent of preferences for one language versus another could be
cooperatively evaluated. However DIAOLS should not be viewed as a
long range solution since it cannot be easily expanded into an
advanced system such as QUL.

2. If the maximum emphasis is to be placed on long term growth
and flexibility of COINS, then we feel that QUL can meet the require-
ments. It is the most advanced and well balanced system we have
surveyed. This approach brings with it at least a three year lapse
that will have to be endured before a network structured around QUL
could exist. During that period we would probably maintain the
present environment.

3. If both long and short range solutions are desirable then
both 1 and 2 above are proposed. In this instance, DIAOLS could
expect an operational life span of about 3 to 6 years; the DIAOLS
life span would not be shortened so much as the introduction of
QUL delayed.

In regard to the long range development of an IS&R facility
within the Agency, we recommend the development of QUL, as it is the
only system we know of with sufficient potential to merit the
expenditure of sizable resources.

DEFENSE INTELLIGENCE AGENCY
WASHINGTON, D. C. 20301

U-4284/MS-3   7 August 1969

MEMORANDUM FOR COINS PROJECT MANAGER

SUBJECT: Network Users Language Specification

1. Reference COINS Master Plan, Part I, FY 1969-1970, dated 2 December 1968.

2. In response to task III. 3. g. in reference, the Network Users Language Specification is submitted for your action as deemed necessary.

25X1

MAJ, USA
Chairman, Computer and Communications
   Interface Panel   1 Enclosure a/s

COMMON COMMUNICATION LANGUAGE (COCOLA)

INTRODUCTION

The ultimate value of the COINS System must be measured in its utility

to those who extract information from it. The users of the COINS System

all require the use of a generalized information management system. The

on-line aspect allows the intelligence analyst to directly interact with

the data bases through the information management system. The analyst must

learn to use the remote terminal and the information management system.

Since there are three information management systems in COINS, the intel-

ligence analyst is faced with learning three separate languages in order to

accomplish the same thing, using different data bases. In view of the fact

the analyst has as his primary job the analysis of intelligence, learning

information management systems must be minimized.

The existence of many systems to perform the same task is undesirable

for a number of reasons. The most significant of these perhaps is the

training required of the users of the systems. An information manager may

have to learn several languages in order to be able to have access to the

data he requires. The information manager, in addition, is denied movement

from one place to another due to different information management systems,

even though he could very well manage the data. Thus, progress and

diversification in his trade is measured in his knowledge of different

information management systems of the same degree of complexity and not

his degree of knowledge of information management.

The information manager uses the notations of the information management system and cares very little about how the computer functions. That is as it should be, the computer being used as the tool under the control of the individual who needs its service. This is made possible through the use of an information management system.

BASIC CONCEPTS

The notation system that is proposed is not to substitute for the notation system of any existing system but rather is designed to serve the information manager and to be converted by computer into the present notation systems. If the notation system gains wide spread use, future information management systems could use it to avoid the extra conversion step which admittedly is time consuming.

The information management systems - TORQUE, FFS, and TILE - evolved over time to serve the needs of the developing agency. Since these systems were developed to process information, they have certain basic similarities.

In order to solve this problem we can start with their similarities and continue to work until we have a notation system which will enable the information manager to take full advantage of all three information management systems.

The information manager formulates his query in the new language. The query is transferred to the retrieval computer, there it must be translated into the notation system of the resident information management system.

In this paper, the notation system will in general deal with functions which are common to all systems. This is the starting point. The area search feature described toward the end is a departure into notating functions that are not universal.

Next 4 Page(s) In Document Exempt